Universal Exact Compression of Differentially Private Mechanisms

Yanxiao Liu, Wei-Ning Chen, Ayfer Özgür and Cheuk Ting Li

NeurIPS 2024

Overview

Background

- In modern data science, large amounts of high-quality data are generated with personal information, which are susceptible to privacy breaches.
- Differential privacy (Warner (1965); Dwork et al. (2006)) is a powerful tool for safeguarding users' privacy by properly randomizing the local data.
- Apart from privacy, communication (of high-dimensional data) often becomes a bottleneck in the system pipeline.

Objective

We intend to answer the following fundamental question: how can we efficiently communicate privatized data?

Kasiviswanathan, S. P., Lee, H. K., Nissim, K., Raskhodnikova, S., & Smith, A. (2011). What can we learn privately?. SIAM Journal on Computing, 40(3), 793-826.

Dwork, C., McSherry, F., Nissim, K., & Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In Theory of Cryptography: Third Theory of Cryptography Conference, TCC 2006, New York, NY, USA, March 4-7, 2006. Proceedings 3 (pp. 265-284). Springer Berlin Heidelberg.

Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. Journal of the American statistical association, 60(309), 63-69.

Related Works

Compression of Differential Privacy (DP) Mechanisms

To compress ϵ-DP mechanisms:

- For $\epsilon \leq 1$, Bassily and Smith (2015) showed that a single bit can simulate any local DP randomizer with a small degradation of utility.
- Bun et al. (2019) proposed a rejection-sampling-based compression technique, which compresses an ϵ-DP mechanism into a 10ϵ-DP mechanism.
- Feldman and Talwar (2021) proposed a distributed simulation approach using rejection sampling with shared randomness.
- In Triastcyn et al. (2021); Shah et al. (2022), importance sampling (or more specifically, minimum random coding (Havasi et al. (2018))) was utilized.
- All these methods are approximate, i.e., the output distribution is distorted.

Bassily, R., & Smith, A. (2015, June). Local, private, efficient protocols for succinct histograms. In Proceedings of the forty-seventh annual ACM symposium on Theory of computing (pp. 127-135).

Bun, M., Nelson, J., & Stemmer, U. (2019). Heavy hitters and the structure of local privacy. ACM Transactions on Algorithms (TALG).

Feldman, V., & Talwar, K. (2021, July). Lossless compression of efficient private local randomizers. In International Conference on Machine Learning (pp. 3208-3219). PMLR.

Shah, A., Chen, W. N., Balle, J., Kairouz, P., & Theis, L. (2022, May). Optimal compression of locally differentially private mechanisms. In International Conference on Artificial Intelligence and Statistics (pp. 7680-7723). PMLR.

Triastcyn, A., Reisser, M., & Louizos, C. (2021). Dp-rec: Private & communication-efficient federated learning. arXiv:2111.05454.

Havasi, M., Peharz, R., & Hernández-Lobato, J. M. (2018). Minimal random code learning: Getting bits back from compressed model parameters. arXiv preprint arXiv:1810.00440.

Related Works

Channel Simulation

One-shot channel simulation (a lossy compression task) aims to find the minimal needed communication over a noiseless channel to "simulate" a channel $P_{Z|X}$.

Related Works

- By Harsha et al. (2007) and Li and El Gamal (2018), $P_{Z|X}$ can be simulated using $I(X;Z) + O(\log(I(X;Z)))$ bits.
- In Harsha et al. (2007), algorithms based on rejection sampling are proposed.
- Dithered quantization (Ziv (1985)) has been used to simulate an additive noise channel in Agustsson and Theis (2020) for neural compression.
- More applications of channel simulation tools:
  - Neural network compression by Havasi et al. (2018)
  - Image compression via variational autoencoders by Flamich et al. (2020)
  - Diffusion models with perfect realism by Theis et al. (2022)
  - Differentially private federated learning by Shah et al. (2022)

Harsha, P., Jain, R., mathcallester, D., & Radhakrishnan, J. (2007, June). The communication complexity of correlation. In Twenty-Second Annual IEEE Conference on Computational Complexity (CCC'07) (pp. 10-23). IEEE.

Li, C. T., & El Gamal, A. (2018). Strong functional representation lemma and applications to coding theorems. IEEE Transactions on Information Theory.

Ziv, J. (1985). On universal quantization. IEEE Transactions on Information Theory, 31(3), 344-347.

Agustsson, E., & Theis, L. (2020). Universally quantized neural compression. Advances in neural information processing systems, 33, 12367-12376.

Poisson Private Representation (PPR)

Poisson Private Representation: Overview(PPR)

- In this paper, we propose a novel algorithm, called **Poisson private representation** (PPR), that is designed to compress and simulate any local randomizer while ensuring local differential privacy.
- The advantages of our PPR are as follows:
  1. **Universality**: Unlike dithered-quantization-based approaches which can only simulate additive noise mechanisms, PPR can simulate any local or central DP mechanism with discrete or continuous input and output.
  2. **Exactness**: PPR enables exact simulation, ensuring that the reproduced distribution perfectly matches the original one, and hence the compressed sample maintains the same statistical properties.
  3. **Communication efficiency**: PPR compresses the output of any DP mechanism to a size close to the theoretical lower bound $I(X; Z)$.
- PPR is the first universal exact compression method for DP mechanisms with an almost-optimal compression size.
  - The methods by Bassily and Smith (2015); Bun et al. (2019); Feldman and Talwar (2021); Shah et al. (2022) are not exact.
  - The methods by Harsha et al. (2007) and Li and El Gamal (2018) do not guarantee privacy.

Our code can be found in `https://github.com/cheuktingli/PoissonPrivateRepr`

Preliminaries

Definition: Differential Privacy

Given a mechanism $\mathcal{A}$ which induces distribution $P_{Z|X}$ of $Z = \mathcal{A}(X)$, we say that it satisfies (ϵ, δ)-DP if for any neighboring $(x, x') \in \mathcal{N}$ and $\mathcal{S} \subseteq \mathcal{Z}$, it holds that[a]

$$\mathbf{P}(Z \in \mathcal{S} \,|\, X = x) \leq e^{\epsilon} \mathbf{P}(Z \in \mathcal{S} \,|\, X = x') + \delta. \quad (1)$$

[a] If a mechanism satisfies $(\epsilon, 0)$-DP, we simply write it as ϵ-DP.

Definition: Poisson Functional Representation (PFR)

Let $(T_i)_i$ be a Poisson process with rate 1, independent of $Z_i \overset{\text{iid}}{\sim} Q$ for $i = 1, 2, \ldots$. Then $(Z_i, T_i)_i$ is also a Poisson process. Fix any distribution P over $\mathcal{Z}$ that is absolutely continuous with respect to Q. Let

$$\tilde{T}_i := T_i \cdot \left(\frac{\mathrm{d}P}{\mathrm{d}Q}(Z_i) \right)^{-1}. \quad (2)$$

The **Poisson functional representation** by Li and El Gamal (2018) selects $Z = Z_K$ with the smallest associated $\tilde{T}_K$, i.e., let $K := \mathrm{argmin}_i \tilde{T}_i$ and $Z := Z_K$.

Li, C. T., & El Gamal, A. (2018). Strong functional representation lemma and applications to coding theorems. IEEE Transactions on Information Theory.

Poisson Private Representation (PPR)

Poisson Private Representation: Construction

Input: x, (ϵ, δ)-DP mechanism $P_{Z|X}$, reference distribution Q, parameter $\alpha > 1$.

1. Generate shared randomness between user and server $(Z_i)_{i=1,2,\ldots} \stackrel{\text{iid}}{\sim} Q$.
2. The user knows $(Z_i)_i$, x, $P_{Z|X}$ and performs:
   1. Generate the Poisson process $(T_i)_i$ with rate 1.
   2. Compute $\tilde{T}_i := T_i \cdot \left(\frac{dP_{Z|X}(\cdot|x)}{dQ}(Z_i) \right)^{-1}$.
   3. Generate $K \in \mathbb{Z}_+$ with
   $$\mathbf{P}(K = k) = \frac{\tilde{T}_k^{-\alpha}}{\sum_{i=1}^{\infty} \tilde{T}_i^{-\alpha}}. \tag{3}$$
   4. Compress and send K (e.g., by Elias delta code).
3. The server, which observes $(Z_i)_i$ and K, outputs $Z = Z_K$.

Remarks

- While the algorithm requires infinite samples, it can be reparametrized to terminate in finite steps.
- When $\alpha = \infty$, PPR reduces to PFR.

Poisson Private Representation (PPR): Theoretic Guarantee

Proposition: Exactness

The output Z of PPR follows $P_{Z|X}$ exactly.

Theorems: Privacy Guarantee

1. **Theorem 4.5** (ϵ-DP of PPR): If the mechanism $P_{Z|X}$ is ϵ-DP, then PPR $P_{(Z_i)_i,K|X}$ with parameter $\alpha > 1$ is $2\alpha\epsilon$-DP.
2. **Theorem 4.8** (Tighter (ϵ, δ)-DP of PPR): If $P_{Z|X}$ is (ϵ, δ)-DP, then PPR $P_{(Z_i)_i,K|X}$ with parameter $\alpha > 1$ is $(\alpha\epsilon + \tilde{\epsilon}, 2(\delta + \tilde{\delta}))$-DP, for every $\tilde{\epsilon} \in (0, 1]$ and $\tilde{\delta} \in (0, 1/3]$ that satisfy $\alpha \leq e^{-4.2}\tilde{\delta}\tilde{\epsilon}^2/(-\ln \tilde{\delta}) + 1$.

Poisson Private Representation (PPR): Theory

Theorem: Communication efficiency

Theorem 4.3 (Compression size of PPR): For PPR with parameter $\alpha > 1$, message K satisfies

$$\mathbf{E}[\log_2 K] \leq D_{\mathrm{KL}}(P(\cdot|x) \| Q(\cdot)) + \frac{\log_2(3.56)}{\min\{\frac{\alpha-1}{2}, 1\}}.$$

As a result, when the input $X \sim P_X$ is random, we have

$$\mathbf{E}[\log_2 K] \leq I(X; Z) + \frac{\log_2(3.56)}{\min\{\frac{\alpha-1}{2}, 1\}}.$$

Hence, K can be encoded into $I(X; Z) + \log_2(I(X; Z) + 1) + O(1)$ bits, close to the theoretical lower bound $I(X; Z)$.

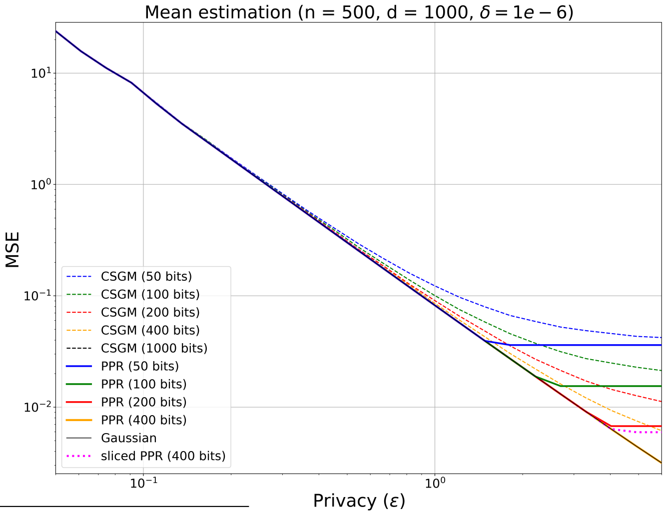
Application: Distributed Mean Estimation

Distributed Mean Estimation

- Consider n users, each with data $X_i \in \mathbb{R}^d$.
- They use **Gaussian mechanism** and send $Z_i \sim \mathcal{N}\left(X_i, \frac{\sigma^2}{n}\mathbb{I}_d\right)$ to server, where $\sigma \geq \frac{C\sqrt{2\ln(1.25/\delta)}}{\epsilon}$. Server estimates the mean as $\hat{\mu}(Z^n) = \frac{1}{n}\sum_i Z_i$.
- Using PPR to compress the Gaussian mechanism:
  - $\hat{\mu}(Z^n) = \frac{1}{n}\sum_i Z_i$ is unbiased, has (ϵ, δ)-central DP.
  - PPR satisfies $(2\alpha\sqrt{n}\epsilon, 2\delta)$-local DP for $\epsilon < \frac{1}{\sqrt{n}}$.
  - The average per-client communication cost is at most $\ell + \log_2(\ell+1) + 2$ bits, where
  $$\ell := \frac{d}{2}\log_2\left(\frac{C^2 n}{d\sigma^2} + 1\right) + \eta_\alpha \;\leq\; \frac{d}{2}\log_2\left(\frac{n\varepsilon^2}{2d\ln(1.25/\delta)} + 1\right) + \eta_\alpha,$$
  and $\eta_\alpha := (\log_2(3.56))/\min\{(\alpha-1)/2,\, 1\}$.
- We compare with Chen et al. (2024) on distributed mean estimation:

Chen, W. N., Song, D., Ozgur, A., & Kairouz, P. (2024). Privacy amplification via compression: Achieving the optimal privacy-accuracy-communication trade-off in distributed mean estimation. Advances in Neural Information Processing Systems, 36.

Application: Distributed Mean Estimation



Chen, W. N., Song, D., Ozgur, A., & Kairouz, P. (2024). Privacy amplification via compression: Achieving the optimal privacy-accuracy-communication trade-off in distributed mean estimation. Advances in Neural Information Processing Systems, 36.

Application: Metric Privacy and Laplace Mechanism

Metric Privacy

For a mechanism $\mathcal{A}$ with $P_{Z|X}$ and a metric $d_{\mathcal{X}}$ over $\mathcal{X}$, it satisfies $\underline{\epsilon \cdot d_{\mathcal{X}}\text{-privacy}}$ (Andrés et al. (2013)) if $\forall x, x' \in \mathcal{X}$, $\mathcal{S} \subseteq \mathcal{Z}$, we have

$$\mathbf{P}(Z \in \mathcal{S} \,|\, X = x) \leq e^{\epsilon \cdot d_{\mathcal{X}}(x,x')} \mathbf{P}(Z \in \mathcal{S} \,|\, X = x').$$

Laplace Mechanism on Geo-indistinguishability

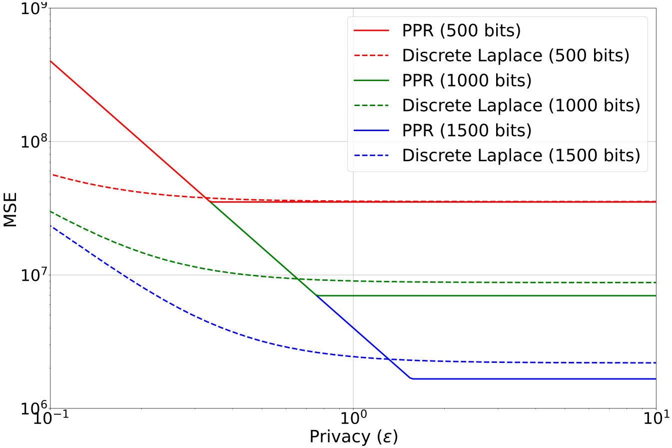
For Laplace mechanism $P_{Z|X}$ with $X \in \{x \in \mathbb{R}^d \big| \|x\|_2 \leq C\}$ and proposal distribution $Q = \mathcal{N}(0, (\frac{C^2}{d} + \frac{d+1}{\epsilon^2})\mathbb{I}_d)$, the output of PPR has MSE $\frac{d(d+1)}{\epsilon^2}$, $2\alpha\epsilon \cdot d_{\mathcal{X}}$-privacy and compression size $\leq \ell + \log_2(\ell + 1) + 2$ bits, where $\ell :=$

$$\frac{d}{2} \log_2 \left(\frac{2}{e} \left(\frac{C^2 \epsilon^2}{d} + d + 1 \right) \right) - \log_2 \frac{\Gamma(d+1)}{\Gamma(\frac{d}{2}+1)} + \frac{\log_2(3.56)}{\min\{\frac{\alpha-1}{2}, 1\}}.$$

We compare with the discrete Laplace mechanism by Andrés et al. (2013).

Andrés, M. E., Bordenabe, N. E., Chatzikokolakis, K., & Palamidessi, C. (2013, November). Geo-indistinguishability: Differential privacy for location-based systems. In Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security (pp. 901-914).

Application: Metric Privacy and Laplace Mechanism



Andrés, M. E., Bordenabe, N. E., Chatzikokolakis, K., & Palamidessi, C. (2013, November). Geo-indistinguishability: Differential privacy for location-based systems. In Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security (pp. 901-914).

Summary

Summary

- We proposed a novel scheme for compressing DP mechanisms, called **Poisson private representation** (PPR).
- Unlike previous schemes which are either constrained on special classes of DP mechanisms or introducing additional distortions on the output, our scheme can compress and **exactly** simulate **arbitrary** mechanisms while providing privacy guarantees.
- PPR provides a compression size that is close to the theoretic lower bound.

Future Works

- Reduce the running time of PPR under certain restrictions. For example, for unimodal $P_{Z|X}$, techniques utilized by Flamich et al. (2022); Flamich (2024) could be useful.

Flamich, G. (2024). Greedy Poisson rejection sampling. Advances in Neural Information Processing Systems, 36.

Flamich, G., Markou, S., & Hernández-Lobato, J. M. (2022, June). Fast relative entropy coding with a* coding. In International Conference on Machine Learning (pp. 6548-6577). PMLR.

Acknowledgement

Yanxiao Liu was partially supported by the CUHK PhD International Mobility for Partnerships and Collaborations Award 2023-24 and CUHK Postgraduate Scholarship. Wei-Ning Chen and Ayfer Özgür were supported by the NSF grant CIF-2213223. Cheuk Ting Li was partially supported by two grants from the Research Grants Council of the Hong Kong Special Administrative Region, China [Project No.s: CUHK 24205621 (ECS), CUHK 14209823 (GRF)].

References

Agustsson, E. and Theis, L. (2020). Universally quantized neural compression. *Advances in neural information processing systems*, 33:12367–12376.

Andrés, M. E., Bordenabe, N. E., Chatzikokolakis, K., and Palamidessi, C. (2013). Geo-indistinguishability: Differential privacy for location-based systems. In *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*, pages 901–914.

Bassily, R. and Smith, A. (2015). Local, private, efficient protocols for succinct histograms. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 127–135.

Bun, M., Nelson, J., and Stemmer, U. (2019). Heavy hitters and the structure of local privacy. *ACM Transactions on Algorithms (TALG)*, 15(4):1–40.

Chen, W.-N., Song, D., Ozgur, A., and Kairouz, P. (2024). Privacy amplification via compression: Achieving the optimal privacy-accuracy-communication trade-off in distributed mean estimation. *Advances in Neural Information Processing Systems*, 36.

Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*, pages 265–284. Springer.

Feldman, V. and Talwar, K. (2021). Lossless compression of efficient private local randomizers. In *International Conference on Machine Learning*, pages 3208–3219. PMLR.

Flamich, G. (2024). Greedy poisson rejection sampling. *Advances in Neural Information Processing Systems*, 36.

Flamich, G., Havasi, M., and Hernández-Lobato, J. M. (2020). Compressing images by encoding their latent representations with relative entropy coding. *Advances in Neural Information Processing Systems*, 33:16131–16141.

Flamich, G., Markou, S., and Hernández-Lobato, J. M. (2022). Fast relative entropy coding with a* coding. In *International Conference on Machine Learning*, pages 6548–6577. PMLR.

Harsha, P., Jain, R., McAllester, D., and Radhakrishnan, J. (2007). The communication complexity of correlation. In *Twenty-Second Annual IEEE Conference on Computational Complexity (CCC'07)*, pages 10–23. IEEE.

Havasi, M., Peharz, R., and Hernández-Lobato, J. M. (2018). Minimal random code learning: Getting bits back from compressed model parameters. *arXiv preprint arXiv:1810.00440*.

Li, C. T. and El Gamal, A. (2018). Strong functional representation lemma and applications to coding theorems. *IEEE Transactions on Information Theory*, 64(11):6967–6978.

Shah, A., Chen, W.-N., Balle, J., Kairouz, P., and Theis, L. (2022). Optimal compression of locally differentially private mechanisms. In *International Conference on Artificial Intelligence and Statistics*, pages 7680–7723. PMLR.

Theis, L., Salimans, T., Hoffman, M. D., and Mentzer, F. (2022). Lossy compression with gaussian diffusion. *arXiv preprint arXiv:2206.08889*.

Triastcyn, A., Reisser, M., and Louizos, C. (2021). Dp-rec: Private & communication-efficient federated learning. *arXiv preprint arXiv:2111.05454*.

Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American statistical association*, 60(309):63–69.

Ziv, J. (1985). On universal quantization. *IEEE Transactions on Information Theory*, 31(3):344–347.